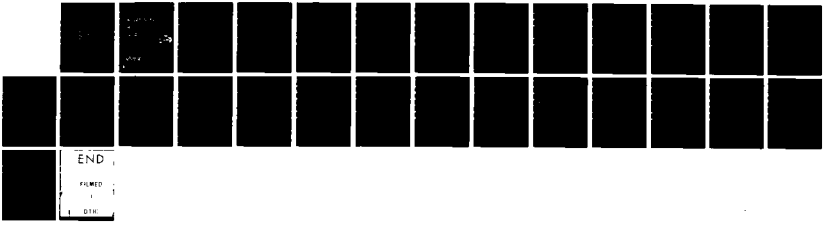
AD-A123 388 DERIVATION AND IMPLEMENTATION OF EXPONENTIAL FUNCTIONS TO MODEL AXIAL SCA.. (U) DIGITAL PROGRAMMING SERVICES INC WALTHAM MASS L E BELSKY ET AL. 30 SEP 82 1/1

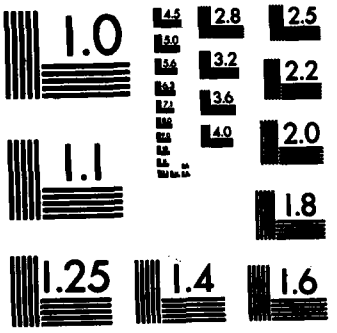
UNCLASSIFIED AFGL-TR-82-0309 F19628-81-C-0141 F/G 12/1 NL



END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

12

AFGL-TR-82-0309

DERIVATION AND IMPLEMENTATION OF EXPONENTIAL FUNCTIONS TO MODEL AXIAL SCATTERING SPECTROMETER DATA DISTRIBUTIONS

LAWRENCE E. SELBY
JAMES P. LALLY
PETER L. MILLER

DIGITAL PROGRAMMING SERVICES, INC.
60 HICKORY DRIVE
WALTHAM, MASSACHUSETTS 02154

30 SEPTEMBER 1982

SCIENTIFIC REPORT NO. 1

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AIR FORCE GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AFB, MASSACHUSETTS 10731

AD A123388

DTIC FILE COPY

DTIC ELECTE

83 01 14 004

Qualified requestors may obtain additional copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFGL-TR-82-0309	2. GOVT ACCESSION NO. A123388	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Derivation and Implementation of Exponential Functions to Model Axial Scattering Spectrometer Data Distributions		5. TYPE OF REPORT & PERIOD COVERED Scientific Report No. 1 1 AUG 81 - 30 SEPT 82 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Lawrence E. Belsky James P. Lally Peter L. Miller		8. CONTRACT OR GRANT NUMBER(s) AF19628-81-C-0141
9. PERFORMING ORGANIZATION NAME AND ADDRESS Digital Programming Services, Inc. 60 Hickory Drive Waltham, MA 02154		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62101F 2310G5BC
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Geophysics Laboratory Hanscom AFB, Massachusetts 01731 Monitor/Barbara A. Main/LYC		12. REPORT DATE 30 SEPT 82 13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

cloud droplet spectra data, distribution functions, mathematical derivation, maximum likelihood method, method of least squares, Newton-Raphson method

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report details the mathematical derivation of the Weibull and gamma distribution functions for the purpose of approximating cloud droplet spectra data via a mathematical model, and analyzes the ability of each function to model the data distributions.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Acknowledgements

The authors wish to thank Ms. Susan Sadofsky for her technical review. Special thanks is given to Mr. Morton Glass of AFGL for his suggestion of the work, for his guidance and for his many constructive criticisms.

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A

TABLE OF CONTENTS

I. INTRODUCTION 5

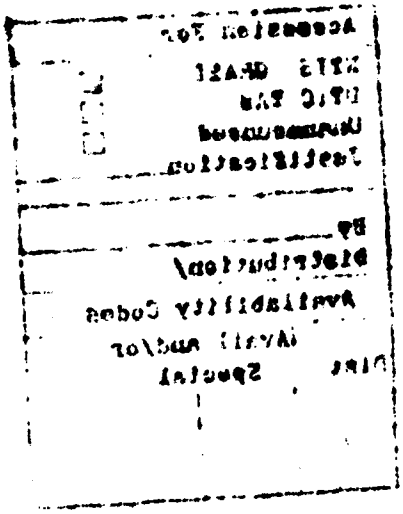
II. WEIBULL DISTRIBUTION FUNCTION 7

III. THE KHRGIAN-MAZIN FORM OF THE GAMMA DISTRIBUTION FUNCTION 14

IV. ANALYSIS 17

V. CONCLUSIONS 24

VI. REFERENCES 25



I. Introduction

In the course of analyzing data under contract AF19628-81-C-0141, an attempt was made to fit certain exponential functions to number density data obtained from axial scattering spectrometer probes.

These data were collected in supercooled clouds associated with aircraft icing. It was noted that the distribution of the number density spectra in the size range of 2 to 30 microns diameter seemed to be describable by the Weibull function or the Khrgian-Mazin[1] form of the gamma distribution. The Weibull function has been used in time-to-failure probability analysis. These functions follow:

Weibull Distribution

$$f(x) = \begin{cases} Nabx^{b-1} e^{-ax^b} & x \geq 0 \quad a>0,\ b>0 \\ 0 & x<0 \end{cases} \tag{1}$$

where N is the distribution function

Further discussion of this Weibull function can be found in Bethea et al.[2]

Gamma Distribution of degree 2 (Khrgian-Mazin)[1]

$$f(x) = \begin{cases} N\frac{a^3}{2} x^2 e^{-ax} & x \geq 0 \quad a>0 \\ 0 & x<0 \end{cases} \tag{2}$$

where N is the distribution function

The method of least squares analysis can be used to determine values of a and b in these functions. It is noted that this effort will require extensive use of numerical analysis to solve the resulting simultaneous non-linear equations. In order to save computation costs it was decided to utilize the maximum likelihood method (MLM) which has the advantage of reducting most solutions to a closed form. The maximum likelihood method (MLM) is described by Breiman[3].

II. Weibull Distribution Function

In order to fit a function, f(x), to a data distribution (y), one mathematical method which can be used is the MLM. This method requires that, for the entire data distribution, the sum of the products of each observed value and the corresponding value of the function

$$\sum_{i=1}^{n} y_i \, f(x_i)$$

must be maximized.

Now, from equation 1, the Weibull is defined as:

$$f(x) = Nabx^{b-1}e^{-ax^b}$$

In order to fit this function to (x_i, y_i), where,

x_i = channel i diameter
y_i = channel i normalized number density

"a" and "b" must be solved. Since the $\ln[f(x)]$ and $f(x)$ are maximized for the same values of "a" and "b", and since $f(x)$ is a product of functions of "a" and "b", it is easier to solve for those values of "a" and "b" which maximize $\ln[f(x)]$.

Therefore, since the MLM implies that the partial derivatives equal zero, "a" and "b" can be solved for by the following equations:

$$\frac{\partial}{\partial a} \sum_{i=1}^{n} y_i \ln(Nabx_i^{b-1} e^{-ax_i^b}) = 0 \qquad (3)$$

and

$$\frac{\partial}{\partial b} \sum_{i=1}^{n} y_i \ln(Nabx_i^{b-1} e^{-ax_i^b}) = 0 \qquad (4)$$

Solving equation (1):

$$\frac{\partial}{\partial a} \sum_{i=1}^{n} y_i \ln(Nabx_i^{b-1} e^{-ax_i^b}) = 0$$

$$\sum_{i=1}^{n} (y_i) \frac{\partial}{\partial a} (\ln(Nabx_i^{b-1} e^{-ax_i^b})) = 0$$

$$\sum_{i=1}^{n} (y_i) \frac{\partial}{\partial a} (\ln N + \ln a + \ln b + (b-1)\ln x_i - ax_i^b) = 0$$

$$\sum_{i=1}^{n} y_i \left(\frac{1}{a} - x_i^b\right) = 0$$

$$\sum_{i=1}^{n} y_i \left(\frac{1}{a}\right) = \sum_{i=1}^{n} y_i x_i^b$$

$$\frac{1}{a} \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} y_i x_i^b$$

Solving for "a" yields:

$$a = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} y_i x_i^b} \tag{5}$$

Similarly solving equation (4):

$$\frac{\partial}{\partial b} \sum_{i=1}^{n} y_i \ln(Nabx_i^{b-1}e^{-ax_i^b}) = 0$$

$$\sum_{i=1}^{n} (y_i) \frac{\partial}{\partial b} (\ln(Nabx_i^{b-1}e^{-ax_i^b})) = 0$$

$$\sum_{i=1}^{n} (y_i) \frac{\partial}{\partial b} (\ln N+\ln a+\ln b+(b-1)\ln x_i-ax_i^b) = 0$$

$$\sum_{i=1}^{n} (y_i) (\frac{1}{b}+\ln x_i-ax_i^b \ln x_i) = 0$$

$$\frac{1}{b} \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} y_i \ln x_i - a \sum_{i=1}^{n} y_i x_i^b \ln x_i = 0 \tag{6}$$

Now, substituting for "a" (from equation (5)) in equation (6) and realizing that $\sum_{i=1}^{n} y_i \neq 0$, yields:

$$\frac{1}{b} + \frac{\sum_{i=1}^{n} y_i \ln x_i}{\sum_{i=1}^{n} y_i} - \frac{\sum_{i=1}^{n} y_i x_i^b \ln x_i}{\sum_{i=1}^{n} y_i x_i^b} = 0 \qquad (7)$$

Since b cannot be solved for in closed form, a numerical method must be used. One such method is the Newton-Raphson method as described by Scarborough[4]. In applying this method, let equation (7) equal g(b). Since the expression

$$\frac{\sum_{i=1}^{n} y_i \ln x_i}{\sum_{i=1}^{n} y_i}$$

is a constant (K) for a given distribution, and independent of the choice of a, b:

$$g(b) = \frac{1}{b} + K - \frac{\sum_{i=1}^{n} y_i x_i^b \ln x_i}{\sum_{i=1}^{n} y_i x_i^b}$$

Now, to solve for the root, b, such that g(b) = 0, an initial guess for $g(b_o)$ is made. Next, solve for b_{n+1} where:

$$b_{n+1} = b_n - \frac{g(b_n)}{g'(b_n)} \quad \text{and where} \qquad (8)$$

$|b_{n+1} - b_n| < \varepsilon$. The object now is to define ε small enough

so that $b \doteq b_{n+1}$. Since the Newton-Raphson method is repetitive by nature, $g'(b_n)$ can be approximated without significant loss of accuracy:

$$g'(b_n) = \lim_{E \to 0} \frac{g(b_n+E)-g(b_n)}{E}$$

Therefore, substituting this approximation for $g'(b_n)$ in equation 8:

$$b_{n+1} = b_n - \frac{g(b_n)E}{g(b_n+E)-g(b_n)}$$

Letting $E = .01(b_n)$ yields:

$$b_{n+1} = b_n - \frac{g(b_n)(.01)(b_n)}{g(b_n+.01b_n)-g(b_n)}$$

$$b_{n+1} = b_n - \frac{.01(b_n)}{\frac{g(1.01b_n)-g(b_n)}{g(b_n)}}$$

or

$$b_{n+1} = b_n - \frac{.01(b_n)}{\frac{g(1.01b_n)}{g(b_n)} - 1} \qquad (9)$$

With "a" and "b" defined, equations (5) and (9) respectively, only the independent factor N remains to be solved.

Now, it can be shown that determining the value of N by way of the MLM, as was done for "a" and "b" in equation 3 and 4, would result in the expression:

$$\frac{1}{N} \sum_{i=1}^{n} y_i = 0$$

Since this expression is undefined for N, another method must be utilized to determine N. Therefore, applying the method of least squares to equation 1 and solving for N using partial derivatives yields:

$$\frac{\partial}{\partial N} \sum_{i=1}^{n} (y_i - Nabx_i^{b-1} e^{-ax_i^b})^2 = 0$$

$$\sum_{i=1}^{n} 2(y_i - Nabx_i^{b-1} e^{-ax_i^b})(-abx_i^{b-1} e^{-ax_i^b}) = 0$$

$$\sum_{i=1}^{n} -2y_i abx_i^{b-1} e^{-ax_i^b} + \sum_{i=1}^{n} 2N(abx_i^{b-1} e^{-ax_i^b})^2 = 0$$

Factoring out 2ab yields:

$$\sum_{i=1}^{n} y_i x_i^{b-1} e^{-ax_i^b} = N(ab) \sum_{i=1}^{n} (x_i^{b-1} e^{-ax_i^b})^2,$$

and solving for N yields:

$$N = \frac{\sum_{i=1}^{n} y_i x_i^{b-1} e^{-ax_i^b}}{ab \sum_{i=1}^{n} (x_i^{b-1} e^{-ax_i^b})^2} \qquad (10)$$

III. The Khrgian-Mazin form of the Gamma Distribution Function

The gamma distribution derivation is accomplished in a manner similar to the Weibull function by fitting f(x) to (x_i, y_i) by way of the MLM. Thus, given that:

$$f(x) = \begin{cases} \frac{Na^3}{2} x_i^2 e^{-ax_i} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad (2)$$

solving for "a" via partial derivatives yields:

$$\frac{\partial}{\partial a} \sum_{i=1}^{n} (y_i \ln(N\frac{a^3}{2} x_i^2 e^{-ax_i}) = 0$$

$$\sum_{i=1}^{n} (y_i \frac{\partial}{\partial a} (\ln N + \ln\frac{a^3}{2} + \ln x_i^2 - ax_i) = 0$$

$$\sum_{i=1}^{n} y_i ((\frac{2}{a^3})(\frac{3a^2}{2}) - x_i) = 0$$

$$\sum_{i=1}^{n} y_i (\frac{3}{a} - x_i) = 0$$

$$\frac{3}{a} \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} y_i x_i = 0$$

Solving for "a" yields:

$$a = \frac{3 \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} y_i x_i} \tag{9}$$

Again, to avoid $\Sigma\ y_i = 0$ the method of least squares analysis is used to solve for N. Thus:

$$\frac{\partial}{\partial N} \sum_{i=1}^{n} (y_i - N(\frac{a^3}{2}) x_i^2 e^{-ax_i})^2 = 0$$

$$\sum_{i=1}^{n} 2(y_i - N(\frac{a^3}{2}) x_i^2 e^{-ax_i})(-\frac{a^3}{2})(x_i^2 e^{-ax_i}) = 0$$

$$\sum_{i=1}^{n} - 2\ y_i\ (\frac{a^3}{2}) x_i^2 e^{-ax_i} + \sum_{i=1}^{n} 2N((\frac{a^3}{2})(x_i^2\ e^{-ax_i}))^2 = 0$$

Factoring out and cancelling $\frac{2a^3}{2}$ yields:

$$\sum_{i=1}^{n} y_i x_i^2 e^{-ax_i} = N\ \frac{a^3}{2} \sum_{i=1}^{n} (x_i^2 e^{-ax_i})^2$$

and solving for N:

$$N = \frac{2 \sum_{i=1}^{n} y_i x_i^2 e^{-ax_i}}{a^3 \sum_{i=1}^{n} (x_i^2 e^{-ax_i})^2} \qquad (12)$$

IV. Analysis

The Weibull and gamma distribution functions were applied to axial scattering spectrometer probe data obtained from 26 discrete data gathering flight passes made by the Air Force Geophysics Laboratory's instrumental C130-E aircraft during the 1979-1980 winter. The instrument probe was designed to collect data in the 2-30 micron range at 2 micron intervals. The resultant 15 data values (channels) for each flight pass comprised the raw data sets. However, due to the excessively noisy data signal present in channel 1, these data were eliminated from the analyses. Thus data channels 2-15 were used as input to the Weibull and gamma functions.

In order to demonstrate the ability of each function to approximate the number density distribution of each flight pass, Figures 1-3 are presented to compare the actual number density distribution for channels 2-15 to the modeled distributions. In addition, the pertinent meteorological data (liquid water content and temperature) for each flight pass are indicated on each figure.

The distributions from the three flight passes illustrated in Figures 1-3 were chosen to typify the range of liquid water content (LWC) and temperature values encountered during the 26 flights. As the figures show, both functions approximate the number density distributions for channels 2-6 with similar accuracy, but the Weibull generally fits the data better than the gamma function for channels 7-9. However, as the number density data levels off for channels 11-15, only the gamma function consistently models this behavior in an acceptable

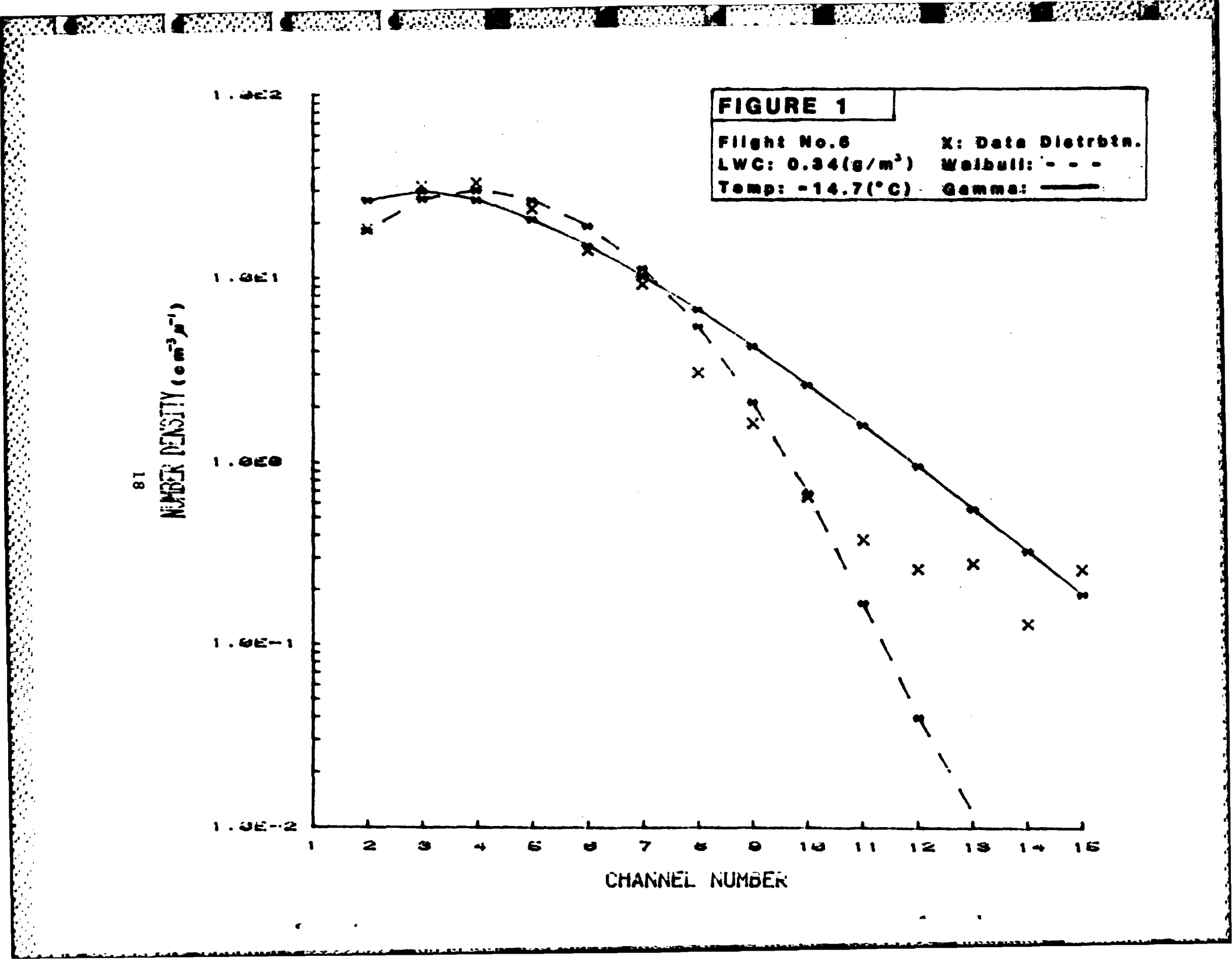


FIGURE 1
Flight No.6
LWC: 0.34(g/m³)
Temp: -14.7(°C)
X: Data Distrbtn.
Weibull: - - -
Gamma: ——
NUMBER DENSITY (cm⁻³μm⁻¹)
1.0E2
1.0E1
1.0E0
1.0E-1
1.0E-2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
CHANNEL NUMBER

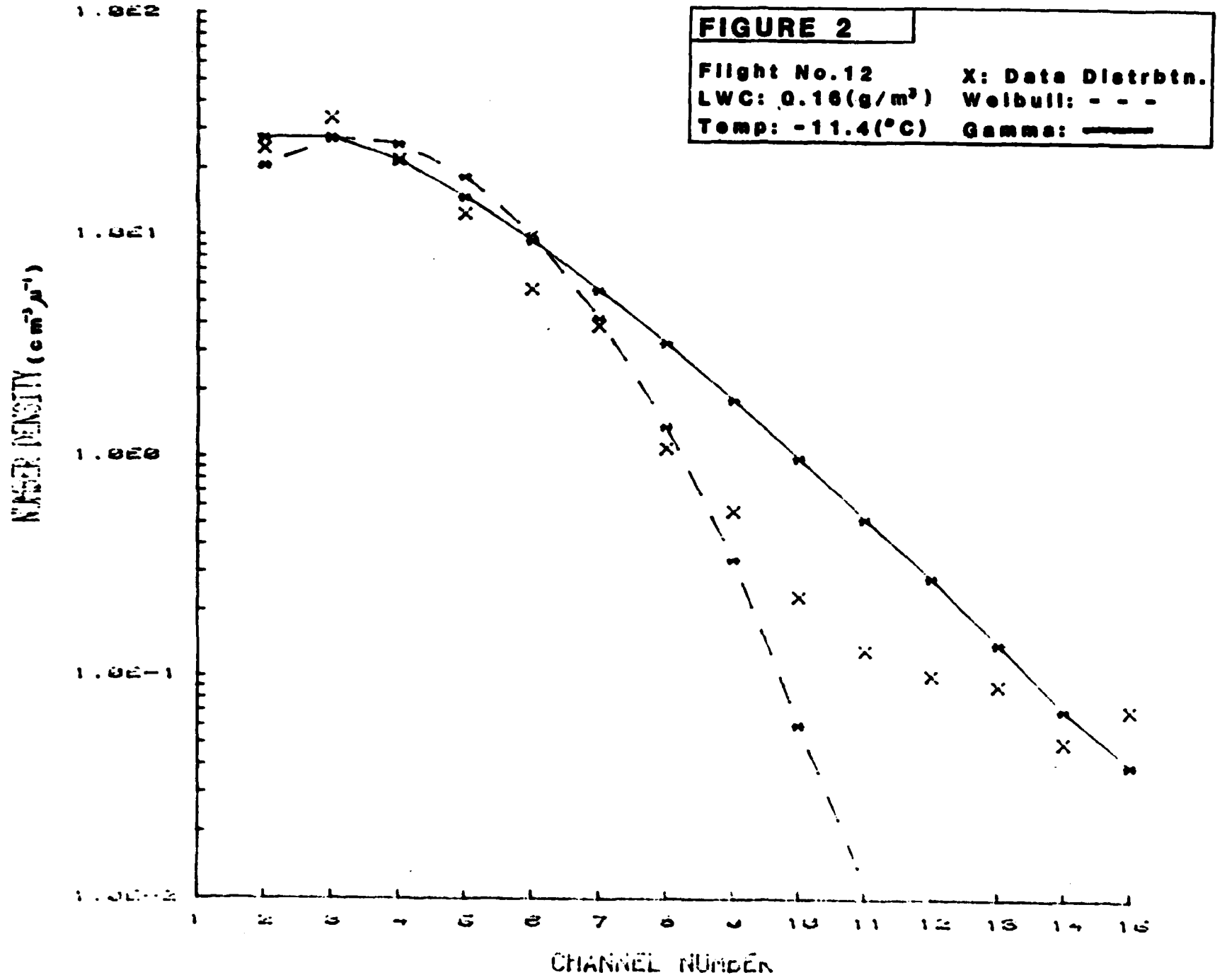


FIGURE 2
Flight No.12
LWC: 0.16(g/m³)
Temp: -11.4(°C)
X: Data Distrbtn.
Weibull: - - -
Gamma: ——
1.0E2
1.0E1
1.0E0
1.0E-1
1.0E-2
NUMBER DENSITY (cm⁻³μm⁻¹)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
CHANNEL NUMBER

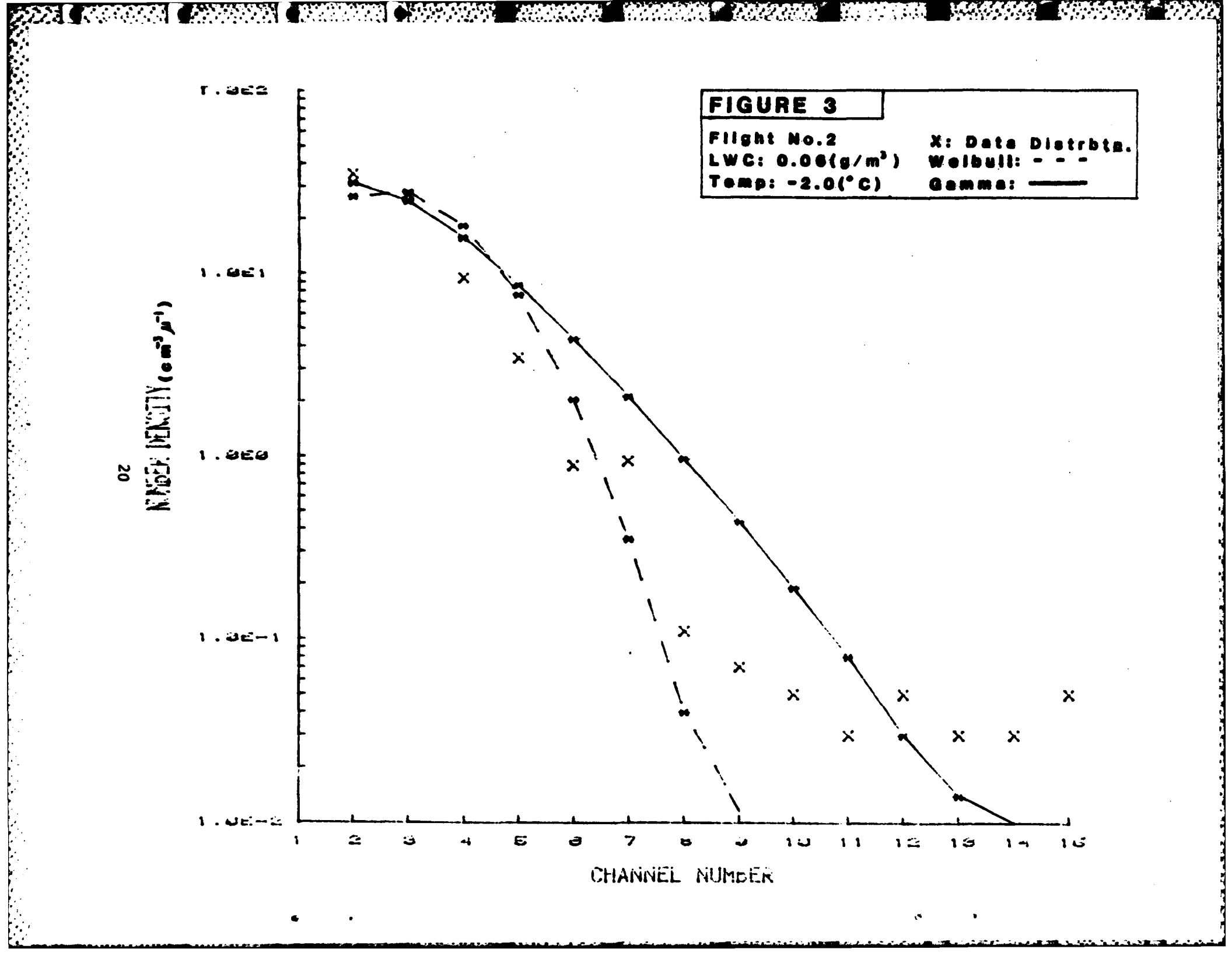


FIGURE 3
Flight No.2
LWC: 0.06(g/m³)
Temp: -2.0(°C)
X: Data Distrbtn.
Weibull: - - -
Gamma: ———
CHANNEL NUMBER

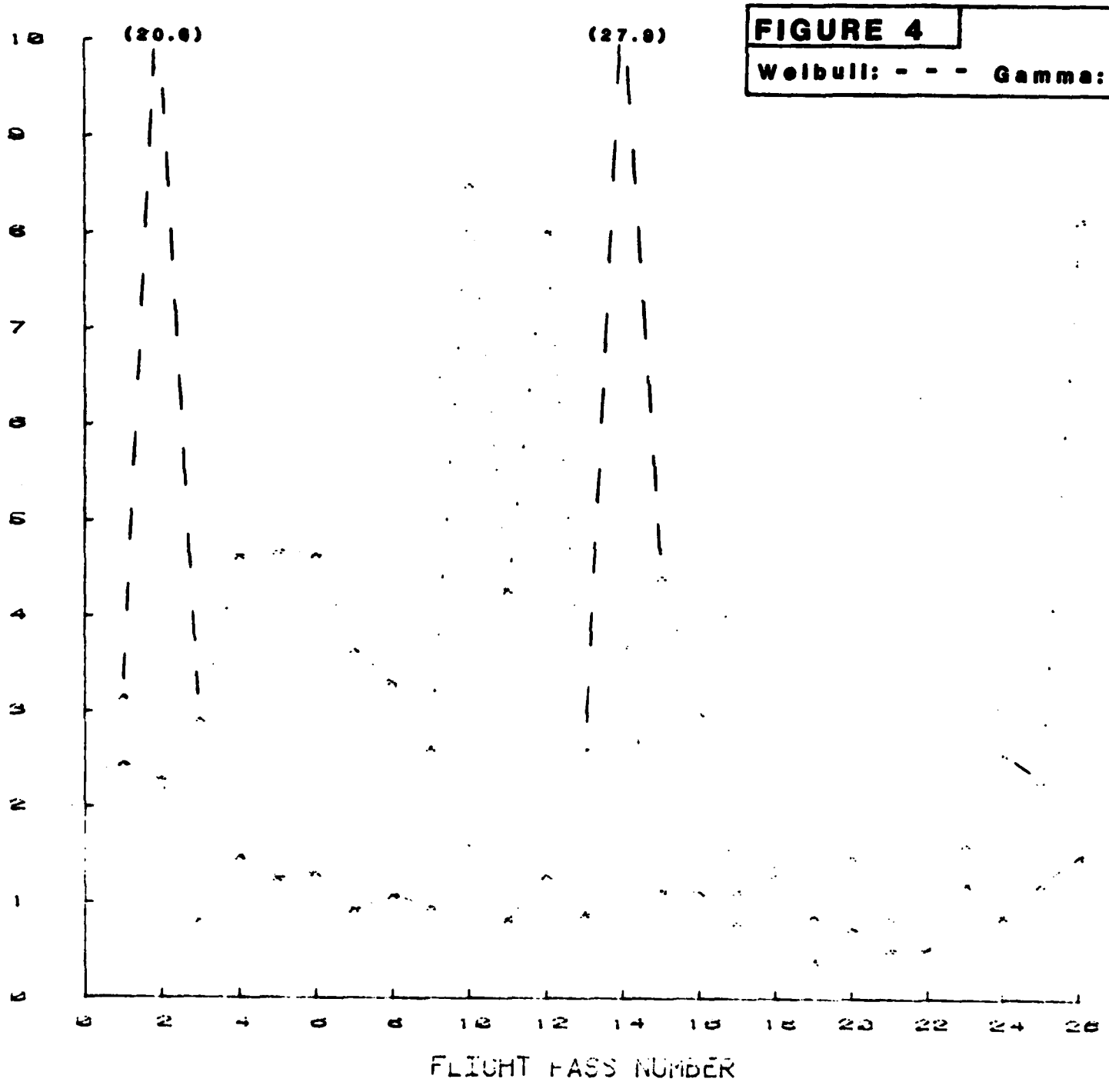


FIGURE 4

Weibull: - - - Gamma: ——

manner. The Weibull's repeated tendency to approach zero at a rapid rate for channels 13-15, 11-15, and 9-15 respectively, necessarily eliminated it as the modeling function of choice. (Note: number density values which are less than the lower limit of the Y axis (1.0E-1) were not plotted.)

The impact of this behavior on the overall performance of the Weibull function is demonstrated in Figure 4, which depicts each functions' log RMS value for each of the 26 flight passes. As can be seen from the figure, 23 of the 26 (88%) gamma RMS values are less than or equal to 2.0 $cm^{-3}\mu^{-1}$; whereas only seven of the 26 (27%) Weibull RMS values are in this category.

This dissimilarity in the performance of the two functions prompted several analyses of the data to determine if some relationship or correlation factor could be discerned between a given data distribution and/or the meteorological conditions (LWC and temperature), during which the data distribution was obtained. Among these analyses was the comparison of the ratio of the number of LWC values above the median LWC value to the number of LWC values below the median for each function. The same comparison of ratios was made using temperature as the variate, with neither comparison evidencing a correlation between the magnitude of LWC or temperature values and the performance of the Weibull function. This lack of correlation between flight pass variables and the Weibull performance held true for log RMS values both above and below the $2.0cm^{-3}\mu^{-1}$ delimiter.

In addition, since the Weibull performed poorest for data channels 11-15 (usually values < 100), an attempt was

made to determine if the Weibull was sensitive to a particular range of data values less than 100. Despite the fact that the Weibull attained log RMS values very close to those of the gamma function for seven flight passes where channels 11-15 were in the 7-70 range (see Figure 4, passes 17-23), there were an equal number of instances where the RMS values deteriorated for flight passes where channels 11-15 were in this same range. (Note: the two flight passes (2 and 14) for which the Weibull attained the worst RMS values were considered an inadequate sample on which to perform analyses. However, it can be mentioned that the gamma function also attained its worst RMS values for the same two passes, albeit an order of magnitude difference exists between these values for the respective functions.)

V. Conclusions

These observations and analyses indicate that for this application, the gamma distribution function is a more accurate and consistent modeling method with which to approximate the data distributions. Subsequence analysis is planned to develop a weighting scheme to refine these modeling techniques.

VI. References

1. Khrgian, A. Kh., Mazin, I.P., Analysis of Methods for Describing Cloud Droplet Distribution Spectra, Trudy TsAO, No. 17, 1956.

2. Bethea, R. M., Duran, B. S., Boullion, T. L., Statistical Methods for Engineers and Scientists, Marcel Dekker, Inc., 1975, p. 66.

3. Breiman, L., Statistics: With a View Toward Applications, Houghton Mifflin, 1973, p. 194 (6.12).

4. Scarborough, J. B., Numerical Mathematical Analysis, Johns Hopkins Press, 1962, p. 199.

2-8
DT